M.R.P. 1036
S.C.II/229
31st January, 1957

AIR MINISTRY
METEOROLOGICAL RESEARCH COMMITTEE

On the attainable accuracy of weather forecasts derived from 24-hr. prebaratic charts

by

R. Murray and C.V. Smith (Meteorological Office)

Summary

The statistical results of an experiment on the 24-hr. forecasting of the weather (at six places scattered over Great Britain) with the aid of 'perfect prebaratics' are presented; for purposes of comparison, similar statistical data from the official press forecasts derived from the official prebaratics and data based on persistence forecasts are also tabulated. The main causes of the larger errors in the forecasts are discussed and illustrative synoptic examples are given.

1. Introduction

A considerable research effort is directed at ways of producing improved forecast pressure (and contour height) charts and the results already published suggest that accurate prebaratics may eventually be produced on a routine basis. It is of interest to consider the probable consequent improvement in the forecasting of the main elements of weather.

In this paper we make an estimate of the success at present of 24-hr. forecasts based on subjectively prepared prebaratics and also of the attainable accuracy of similar 24-hr. forecasts based on perfect prebaratics.

2. Method of obtaining the forecasts

For some years past, the Central Forecasting Office has supplied spot forecasts (i.e., forecasts at a given place and time) for the main elements of weather on the 1200 G.M.T. prebaratic issued to certain newspapers. The aim is to cover those broad features of the weather in which the public are chiefly interested and no attempt is made to give fine detail. The forecasts are of wind, temperature and precipitation (or the state of the sky if no precipitation is foreshadowed) and the expected conditions are summarised with the aid of the Beaufort letters b, bc, c, etc. These forecasts were used to estimate current forecasting success; a comparable set of forecasts from perfect prebaratics was also prepared. Perfect prebaratics were given by copies of the analyses of subsequent 1200 G.M.T. working charts omitting all the observations.

It was decided to prepare at the same time an additional set of forecasts from the official prebaratics as a control: comparison with the official forecasts would enable an estimate to be made of the probable success with perfect prebaratics under normal working conditions.

/Period

Period forecasts of a particular type were also attempted and covered the six-hour interval centred on the prebaratic time, i.e., the period 0900 to 1500 G.M.T.: forecasts were prepared on the basis of both the official and the perfect prebaratics.

To simulate working conditions in the forecast room it would have been necessary to set up all surface and upper air charts, tephigrams and auxiliary aids which the forecaster employs up to the time of preparation of the 1200 G.M.T. prebaratic. Such a thorough-going experiment was not possible and display charts were limited to the surface charts up to 1200 G.M.T., the 1000-500mb. thickness charts up to 0300 G.M.T., charts showing the rainfall and minimum temperature measured at 0900 G.M.T. and the rainfall and maximum temperature read at 2100 G.M.T. the previous day. Charts of normal air and sea temperatures were also to hand.

100 individual days spread over the two years 1951 and 1952 were used for the forecasts. The occasions were selected at random by two members of the Forecasting Research Division, the only proviso being that no two consecutive days should be chosen. Generally several days separated each occasion and four or five forecast days were taken from each of the 24 months; this ensured that the occasions were reasonably independent.

On each of the 100 occasions, forecasts were made for six places, namely Stornoway (Isle of Lewis), Dyce (Aberdeen), Renfrew (Glasgow), Ringway (Manchester), London Airport (London) and Pembroke Dock (southwest Wales). The official press forecasts are prepared for these places and meteorological observations are made at each.

In extracting the data from the official press forecasts, after the experimental forecasts had been completed, it was unfortunately impossible to trace the press forecasts for 12 days (i.e., 72 spot forecasts) so that the total number of spot forecasts is 528.

3. Marking the forecasts

(a) The marking system.

The system we have adopted is simple and objective, though arbitrary. The same system is employed for checking both spot and period forecasts.

Forecasts were assessed on a point scale 0 to 5 according to the degree of error, the top score being 5 for a correct forecast. The basis of the marking system is given in Fig.1, where the normal Beaufort letters are used. Briefly the points to be deducted from the maximum possible score of 5 are equal to the minimum number of steps or dashes separating the weather elements as forecast and observed. Hail showers (ph) and thunder (tl) are not included in the scheme shown in Fig.1 but their treatment is indicated in Table I where the complete scoring system is given.

(b) Summarising the weather: spot forecasts.

The observed weather was in all cases that recorded in the Daily Pocket Registers of the six stations.

The aim in the spot forecasts was to predict the weather at 1200 G.M.T. Since the forecasting of showers or indeed of any type of precipitation at a fixed time and place is difficult it was considered best to compare the forecasts with 'present weather' and 'weather in the past hour'. If no precipitation occurred at 1200 G.M.T. or in the past hour the state of the sky at 1200 G.M.T. was taken without considering the variation of cloud in the past hour. Total cloud amount was used and the symbols b, bc and c were taken as 0-2, 3-6, and 7-8 octas respectively. If precipitation occurred at 1200 G.M.T. or in the past hour, the observed weather was precipitation of the type reported. In a few cases different types of precipitation

/occurred

occurred at 1200 G.M.T. and in the past hour and in such cases the procedure (laid down beforehand) was as follows:- Thunder takes precedence; general precipitation (r, rs, s) takes precedence over showers (pr, prs, ps, ph); assessment is r if r and rs occur (similarly for shower types); assessment is rs if r and s or r, rs and s occur (similarly for showers); assessment is s if rs and s occurs (similarly for showers).

(c) Summarising the weather: period forecasts.

In the six-hour interval of the period forecasts quite large changes of weather may arise and a forecast procedure which could cover a sequence of events was obviously desirable. In the event, however, simplified forecasts of the dominant weather element of the period were prepared in terms of one of the symbols b, bc, c, etc. The symbol to be allocated to the observed weather of the period was determined as follows:-

(i) No precipitation observed.

The weather types b, bc, c and f, were weighted by the factors 1, 2, 3 and 4 respectively and the number of each type occurring in the seven hourly observations was noted. The characteristic weather was taken as the weighted mean of the seven observations. For example, if the sequence of observations is f, b, bc, bc, bc, bc, c the weighted mean is $\frac{1x1 + 4x2 + 1x3 + 1x4}{7}$ or 2 to the nearest whole number, bc was therefore taken as the characteristic weather of the period. One exception was made to this rigorous treatment; if f occurred at four or more hours it was taken as characteristic of the period.

(ii) Precipitation observed.

If any of the precipitation elements were reported at least once between 0900 and 1500 G.M.T., the characteristic weather was taken as precipitation. When more than one type of precipitation was reported during the period, the rules of precedence set out for the spot forecast in section 3(b) were adopted.

In making the period forecasts, the characteristic weather was estimated with these rules of precedence in mind.

4. Accuracy of spot forecasts of weather

Statistical tests of significance have not been applied to the results given below since to have done this would have entailed considerable labour in working out the cross-correlations between the weather at the six points for both types of forecast and the correlations between the forecasting methods, as well as a very careful examination of the theoretical basis for the application of the known statistical tests of significance to such data. The resulting major statistical investigation would have been scientifically interesting but it was not possible for the authors to undertake such work.

(a) Accuracy of the individual forecasts.

Let the press forecasts on the official prebaratics be denoted by F_1, the control forecasts on the official prebaratics by F_2, the forecasts on the basis of persistence by F_3, the forecasts on the perfect prebaratics by F_4 and the probable perfect prebaratic forecasts under normal working conditions by F_5.

The contingency Tables II and III indicate how the predicted weather was related to the observed weather in the press forecasts F_1 and the perfect prebaratic forecasts F_4, respectively. Tables II and III show that, broadly speaking, the forecasts of precipitation

are substantially more accurate in F_4 than in F_1, but that there is little to choose between the forecasts of dry weather. In F_4 there was a surprisingly marked reluctance to forecast b; this is probably a reflection of personal caution in forecasting and not primarily an attribute of the perfect prebaratic. The sample of forecasts is much too small for any firm conclusions to be drawn about the success of forecasting many of the individual elements, but probably adequate to obtain the overall average success in forecasting with the aid of the two types of prebaratic.

(b) The overall average scores.

Table IV shows the overall average score (per cent.) and the percentage success in the various types of forecast F_1 to F_5. There is a small difference between the official press forecasts (F_1) and the control forecasts (F_2). This difference has been employed to estimate the success probable when using perfect prebaratics under normal working conditions. The estimated average score of F_5 is (average score of F_4 - average score of F_2 + average score of F_1): the estimated m.s. error of F_5 is (m.s. error of F_4 - m.s. error of F_2 + m.s. error of F_1).

It is not strictly correct to compute percentage success since the score parameter is not a continuous variable, but nevertheless a rough idea of the success in forecasting the variability is given by the statistic

$$100\left(1 - \frac{\text{m.s. error in score of Fn}}{\text{m.s. error in score of } F_3}\right)^{\frac{1}{2}}.$$ (Priestley (1)) On this basis the press forecasts have a success of about 50 per cent. and the success with the perfect prebaratic forecasts under normal working conditions is likely to be nearly 75 per cent.

From Table IV there is little doubt that the use of perfect prebaratics substantially improves the accuracy of the forecasts, but quite a large residual error still remains.

(c) Daily scores.

The accuracy in forecasting for Great Britain as a whole may be viewed in a rather different manner by treating the total score achieved on each day as the individual unit. Thus wholly accurate forecasts score 6 x 5 = 30 points. Table V gives the frequency of occurrence of various ranges of scores for the several sets of forecasts F_1 to F_5.

It is fairly clear that the forecasts of F_1 are slightly better than those of F_2 at the top and a little inferior at the bottom of the scale. It is rather difficult to allow precisely for the difference between F_1 and F_2 in modifying F_4 to obtain F_5, but the estimate given for F_5 is not likely to be far from the truth. F_5 is significantly better than F_1 and both are appreciably more accurate than F_3.

A legitimate question to ask is how often the perfect prebaratic forecasts are significantly better than the press forecasts.

To answer this question the differences between the daily scores of F_4 and F_1 were noted and the frequencies of various differences are shown in Table VI. As only 88 days were available for the press forecasts, the missing 12 days were taken as the differences between the daily scores of F_4 and F_2.

Table VI shows that on 68 days the perfect prebaratic forecasts were better than the press forecasts, on 29 occasions they were worse and on the remaining 3 days the scores were identical. If the

/assumption

assumption is made that in practice there is no significant difference between the two sets of forecasts unless the daily scores differ by more than 5 points, then on 32 days the perfect prebaratic forecasts were significantly better than the press forecasts, on 9 days they were worse and on 59 days the scores were not significantly different. The fact that on 9 per cent. of occasions the official prebaratics, which were always in error to some extent, should produce better forecasts than the accurate prebaratics emphasises that the prebaratic chart does not uniquely determine the weather in the present state of our knowledge.

If it were possible to tabulate the difference between the daily scores of F_5 and F_1 the figures given in Table VI would no doubt require slight adjustment, but the broad pattern would probably be the same. We may guess that about 35 per cent. of the F_5 forecasts would be significantly better and some 5 per cent. worse than, the F_1 forecasts (significance implying that the daily scores must differ by more than 5 points).

(d) Forecasts of precipitation or no precipitation.

So far we have discussed the overall forecasting success in terms of our arbitrary scoring system, but it is of some importance to access the accuracy of forecasting precipitation. Tables VII and VIII (which are readily obtained from Tables II and III) show the success when the forecasting problem is brought down to saying simply whether or not precipitation will occur. Following Crossley (2), the various subsidiary statistics shown in Table IX are readily computed.

In Table IX the figures for F_5 have been obtained by adjusting the figures for F_4 to allow for the small difference between sets F_1 and F_2. According to Crossley (2) forecasting is worthwhile when the index I is greater than 50 per cent. and on this basis all the forecast methods are useful. Persistence plays quite an important part however. For F_1, the forecasting success over and above persistence is only 15 per cent., whereas for F_5 it is estimated as 48 per cent. The improvement of F_5 over F_1 is about 39 per cent.

It will be noted that both F_1 and F_4 achieve considerable success in forecasting dry weather. On the other hand the forecasting of precipitation is rather poor with F_1 but much better with F_4, the improvement being about 45 per cent. The overall accuracy of perfect prebaratic forecasts of this type (black-white) under normal working conditions (F_5) is probably about 80 per cent., compared with 70 per cent. for the overall accuracy of the official press forecasts (F_1).

5. Accuracy of period forecasts of weather

Only a few results are presented and briefly discussed. Let the period forecasts in the official prebaratic be denoted by PF_2, the period forecasts on the basis of persistence by PF_3 and the period forecasts on the perfect prebaratics by PF_4.

Table X shows the overall average success and the percentage success in the various period forecasts PF_2 to PF_4. A direct comparison may be made with Table IV and it is seen that the success with period forecasts is generally less than for the spot forecasts. Apart from the difficulty of forecasting the characteristic weather of the period (see Section 3(c)) the chief reason for the lower success was that precipitation was forecast too often. Nevertheless both PF_2 and PF_4 produce a big improvement over persistence PF_3.

/The

The Crossley statistics discussed in Section 4(d) may be prepared for the period forecasts of precipitation. These are shown in Table XI and a comparison may be made with Table IX. Clearly the accuracy of forecasts of some precipitation during the period 0900 to 1500 G.M.T. is much higher than that of spot forecasts of precipitation; but this result arose largely because precipitation was forecast too often. The period forecasts of no precipitation (PF_2 and PF_4) were correspondingly less successful than the spot forecasts of no precipitation.

6. Results of the temperature forecasts

The average error (without regard to sign) and the root mean square temperature errors of the forecasts F_1 to F_5 are set out in Table XII.

The estimated average error of F_5 is (average error of F_4 - average error of F_2 + average error of F_1): the estimated r.m.s. error of F_5 is (m.s.e. of F_4 - m.s.e. of F_2 + m.s.e. of F_1)$^{\frac{1}{2}}$. Following Priestley (1) the percentage success in forecasting the 24-hour variability is also listed.

As regards the individual stations, the temperature variability was least at Stornoway (r.m.s. 24-hour change = 3.59°F) and greatest at London (r.m.s. 24-hour change = 5.53°F). In most of the forecasts, the size of the temperature errors roughly reflected the temperature variability at each station. It will be observed that the average errors of the F_1 and F_5 forecasts are tolerably small; but persistence forecasts do not produce large errors on the average.

Broadly speaking, errors of less than 5°F in 24-hr. forecasts are regarded as acceptably small and errors equal to or greater than 10°F as serious. Forecasts with errors of intermediate magnitude are generally not valueless. The precise size of the error is not, however, always a criterion of the usefulness of the temperature forecast; an error of a few degrees may be serious at about the freezing point, whereas an error of 10°F may be unimportant when the temperatures are very high. In Table XIII are listed the percentage occurrence of the errors (irrespective of sign) in these broad categories.

The estimated distribution of the errors based on perfect prebaratics under normal working conditions (F_5) is derived on the assumption (a somewhat optimistic one) that the difference between the frequencies of the two largest errors in F_1 and F_2 can be deducted from the corresponding figures in F_4 and the residual percentage added to the frequency of the small error group in F_4 to obtain the figure 85 per cent. for F_5 in Table XIII.

It will be observed that a "no change" forecast (F_3) is good nearly three times out of four. Nevertheless, the press and perfect prebaratic forecasts are significantly superior to the forecasts of no change even though persistence is clearly such an important factor.

The occurrence of large temperature errors (i.e., errors $\geqslant$ 10°F) is particularly interesting. With F_1 there were 78 days (out of 88) when no large error obtained, 6 days when only one station and 4 days when two stations were affected. With F_4, 91 days (out of 100) had no large errors, 8 days had only one station affected and 1 day had three stations affected. No doubt F_5 would have an appreciably better record than F_4 as regards large errors. Thus even on the occasional days when large errors occur it is rare for more than about one third of Britain to be affected.

The large temperature errors generally occurred with F_4 in dry weather. There were 11 cases of errors $\geqslant$ 10°F in 600 individual forecasts. On four of these occasions fog was concerned. On three other occasions the nocturnal temperature change was wrongly estimated and on the remaining four occasions the state of the sky was incorrectly forecast.

/With

With F_1 the reasons for the large temperature errors were much as for F_4, but prebaratic errors were of course added complications and accounted for roughly two-thirds of the occasions of large temperature errors with F_1.

Large positive errors (i.e., forecast temperatures too high by 10°F or more) were nearly twice as frequent as large negative errors in F_4 and this bias was even more pronounced in the case of F_1. The majority of these large positive errors appear to have occurred after a very cold night.

It is clear from Tables XII and XIII that the F_5 forecasts are significantly better than those of F_1. Further improvement can only come from research into the fundamental problems associated with temperature variations and from the study of individual localities. Nevertheless it is scarcely likely that a substantial reduction of the size of the average error shown in row 5 of Table XII will be possible; probably the main result of any investigations will be to minimise the number of the larger errors.

7. Synoptic examples of spot forecasts

The daily scores for F_1 and F_4 have been examined in relation to the synoptic charts and a few illustrative examples are now presented.

An example of an excellent forecast on the perfect prebaratic (F_4 score = 30 points) and a poor press forecast (F_1 score = 10 points) is given in Fig.2 (19 Jan. 1951); it will be noted that the temperature forecasts were both good (i.e., all temperature errors $\leqslant 4°F$). In this case the press forecast was primarily in error because the prebaratic was poor.

However it is possible to have a good F_4 forecast and a rather poor press forecast even when the prebaratic is reasonably good; such an example is shown in Fig. 3 (17 Sept. 1952) where the general pressure distribution is quite well forecast though the error in the position of the front is about 120 miles (F_4 score = 27 points and F_1 score = 18 points). In this example the F_4 temperature forecasts were very good; the F_1 temperature forecasts were also good apart from an error of 8°F at Glasgow.

Fig. 4 (11 May 1951) is an example of an excellent prebaratic; both forecasts were very good (F_4 score = 27 points and F_1 score = 29 points); the temperature forecasts were quite good. When the scores for both F_1 and F_4 were good the synoptic situations were largely anticyclonic (or anticyclonic-block) and usually the prebaratic error was small or unimportant. However, the larger temperature errors tended to occur in such situations.

There were 9 cases when the perfect prebaratic forecasts were poor (F_4 scores = 11 to 15, see Table V). In most of these cases the press forecasts scored significantly more points. The main reason for the low F_4 scores was the difficulty of estimating the distribution of precipitation. In such cases the F_1 scores were higher more or less fortuitously. Figs. 5 (31 Oct. 1952) and 6 (19 Feb. 1951) are good examples. In Fig.5 the F_4 score was only 12 points because the cloud and rainfall associated with the warm front was rather abnormal and incorrectly forecast. The official prebaratic was appreciably in error in placing the front too far to the west; with such a forecast of the frontal position it was reasonable to predict mainly dry weather for Britain as was done both in the press forecasts (F_1 score = 24 points) and in the control forecasts on the official prebaratic (F_2 score = 25 points). Mainly dry weather was actually observed at 1200 G.M.T. over Britain, but it is unlikely that this would have been predicted if the prebaratic had been an accurate one. Incidentally, both temperature forecasts were good. Fig. 6 shows that the official prebaratic was in error in predicting slight anticyclonic curvature of the isobars over Britain, but in the event the isobaric curvature was cyclonic and the pressure was lower than predicted. The actual chart was surely more favourable for showery precipitation than was the official prebaratic, yet 5 of the stations (out of 6) reported dry weather. The scores for F_1, F_2 and F_4 were 24, 24 and 15 points respectively: the temperature forecasts were good.

/The

The last example, Fig. 7 (5 May 1952), is one in which the perfect prebaratic forecast (F_4 score = 12 points) and the press forecast (F_1 score = 14 points) were both poor; the temperature forecasts were good. The precise distribution of precipitation over Britain was clearly difficult to forecast. In cases such as this, a general worded forecast for a large area is sometimes easy to make and may be tolerably accurate.

8. Conclusions

(i) There is no doubt that accurate prebaratics can improve forecasting in the British Isles. For the type of forecast chiefly considered (24-hr. spot forecasts of the dominant weather) the improvement is substantial, even with present knowledge and techniques.

With the objective but arbitrary system adopted here for marking the forecasts, the average score of the official press forecasts (F_1) is 64.8 per cent., of persistence forecasts (F_3) is 55.8 per cent. and of the perfect prebaratic forecasts under normal working conditions is 75.2 per cent.

From the daily forecast scores for Great Britain as a whole, it appears that some 35 per cent. of the F_5 forecasts are likely to be significantly better and some 5 per cent. worse than the F_1 forecasts.

(ii) The overall accuracy in forecasting precipitation or no precipitation is 70, 64 and 80 per cent. for F_1, F_3 and F_5 respectively, whilst the 'index of usefulness' in such forecasting is about 64, 58 and 78 per cent. for F_1, F_3 and F_5 respectively. The improvement of F_5 over F_1 is 39 per cent.

(iii) In 24-hr. temperature forecasts, the average errors are 2.98°F for the official press forecasts (F_1), 3.55°F for the persistence forecasts (F_3) and 2.66°F for the perfect prebaratic forecasts under normal working conditions (F_5): the percentage success in forecasting the 24-hr. variability of temperature is 55 for F_1 and 67 for F_5 (0 for F_3, of course). Large temperature errors, i.e., errors $\geq$ 10°F, occur on 2.6, 5.2 and 1.0 per cent. of occasions in F_1, F_3 and F_5 respectively.

(iv) A large residual error still remains when forecasts of the general nature attempted by C.F.O. are made with accurate prebaratics. Any additional improvement can only come from a clearer understanding of the physical processes of weather. (Accurate prebaratics could possibly effect some further improvement in association with the development of local objective techniques but this question has not been examined.)

(v) Much interest these days centres on "numerical forecasting" where the objective is to arrive at computational methods of predicting the future patterns of air flow. The approximations of the mathematical models employed make it unlikely that accurate prebaratics will be consistently produced. The field of vertical motion at various levels in the atmosphere can be made available and should prove an invaluable aid in the forecasting of precipitation but the methods do not yet locate fronts. From these considerations, it may well be that the accuracy of the resulting weather forecasts will not differ significantly from those of our forecasts based on perfect prebaratics, with fronts marked (but no direct information on the field of vertical motion).

9. Acknowledgements

We should like to thank Mrs. C.V. Smith for assistance with the statistical work and Mr. J.S. Sawyer for his encouragement and criticism.

/REFERENCES

REFERENCES

1. Priestley, C.H.B. Accuracy of upper air forcasts. Meteorological Research Papers, Air Ministry, London, No. 222, 1945.

2. Crossley, A.F. Measures of success in forecasting. Met. Mag., London, 83, 1954, pp. 66 and 139.

Table I

Scoring system for weather forecasts

Observed weather

Forecast weather	b	bc	c	f	pr	prs	ps	ph	tl	r	rs	s
b	5	4	3	3	1	1	1	1	0	0	0	0
bc	4	5	4	3	2	2	2	2	0	0	0	0
c	3	4	5	3	2	2	2	2	0	0	0	0
f	3	3	3	5	0	0	0	0	0	0	0	0
pr	1	2	2	0	5	4	3	4	3	3	2	1
prs	1	2	2	0	4	5	4	4	3	2	3	2
ps	1	2	2	0	3	4	5	4	3	1	2	3
ph	1	2	2	0	4	4	4	5	3	2	2	2
tl	0	0	0	0	3	3	3	3	5	3	3	3
r	0	0	0	0	3	2	1	2	3	5	4	3
rs	0	0	0	0	2	3	2	2	3	4	5	4
s	0	0	0	0	1	2	3	2	3	3	4	5

Table II

Results of press forecasts (F_1)

Observed weather

Forecast weather	b	bc	c	f	pr	prs	ps	ph	tl	r	rs	s	Total
b	18	7	11	3	1					1			41
bc	30	70	60	7	22		1			22			212
c	5	32	44	1	4		1	1		27		1	116
f			5							2			7
pr	3	13	14		18			2		14	1		65
prs		1						1					2
ps						1	2	1					4
ph							1						1
tl													0
r	3	16	23	1	8		2			25	1		79
rs			1										1
s													0
Total	59	139	158	12	53	1	7	5	0	91	2	1	528

Table III

Results of perfect prebaratic forecasts (F_4)

Observed weather

Forecast weather	b	bc	c	f	pr	prs	ps	ph	tl	r	rs	s	Total
b	3	5	2										10
bc	40	69	52	6	11		1			7			186
c	16	52	89	2	9					21			189
f		3	2	2						2			9
pr	2	23	17	1	28			1		13	1	1	87
prs		3	1		2		1	3		1			11
ps					1	1	5						7
ph								1					1
tl													0
r		8	19	1	7					61			96
rs			1							1	1		3
s		1											1
Total	61	164	183	12	58	1	7	5	0	106	2	1	600

Table IV

Overall average score (per cent.) and percentage success in forecasts F_1 to F_5

	Average score	Percentage success	No. of obs.
F_1 press forecasts on official prebaratics	64.8	49.9	528
F_2 control forecasts on official prebaratics	63.6	44.8	600
F_3 persistence forecasts	55.8	0	600
F_4 forecasts on perfect prebaratics	74.0	71.2	600
F_5 probable perfect prebaratic forecasts under normal working conditions	75.2	74.3	-

Table V

Frequency of daily scores for forecasts F_1 to F_5

	Scores						
	30-26	25-21	20-16	15-11	10-6	5-0	No. of obs.
F_1	15	28	24	15	4	2	88
F_2	14	31	29	20	6	0	100
F_3	4	22	34	25	13	2	100
F_4	25	45	21	9	0	0	100
F_5	28	44	19	9	0	0	-

Table VI

Frequency of various differences between the daily scores of perfect prebaratic forecasts (F_4) and press forecasts (F_1)

	Positive differences							Negative differences			
	30-26	25-21	20-16	15-11	10-6	5-1	0	1-5	6-10	11-15	Total
$F_4 - F_1$	1	0	3	7	18	31	3	18	5	2	88
$F_4 - F_1$ (modified)	1	0	3	7	21	36	3	20	6	3	100

Table VII

Press forecasts of precipitation (F_1)

	Forecast	Not forecast	Total
Precipitation	(A) 77	(B) 83	160
No precipitation	(C) 293	(D) 75	368
Total	370	158	528

Table VIII

Forecasts of precipitation from perfect prebaratics (F_4)

	Forecast	Not forecast	Total
Precipitation	(A) 129	(B) 51	180
No precipitation	(C) 343	(D) 77	420
Total	472	128	600

Table IX

Summary of various statistics relating to forecasts of precipitation in the forecasts F_1 to F_5 (per cent.)

	a	c	c'	I	J
F_1	70.0	48.1	79.6	63.9	15.1
F_2	68.5	50.0	76.2	63.1	13.2
F_3	64.0	41.7	73.3	57.5	-
F_4	78.7	71.7	81.7	76.7	45.2
F_5	80	-	-	78	48

overall accuracy of forecasts = $a = \frac{A + C}{A + B + C + D}$;

accuracy of forecasts of precipitation = $c = \frac{A}{A + B}$;

accuracy of forecasts of no precipitation = $c' = \frac{C}{C + D}$;

index of usefulness = $I = \frac{1}{2}(c + c')$
($I \geqslant \frac{1}{2}$ or 50 per cent. for worthwhile forecasting);
index measuring forecasting success over and above persistence forecasting = $J = \frac{I - I_o}{1 - I_o}$ where I_o is index of usefulness on basis of persistence.

Table X

Overall average score and percentage success in period forecasts PF_2 to PF_4

	Average score (per cent.)	Percentage success	No. of obs.
PF_2 period forecasts on official prebaratics	62	49	600
PF_3 period persistence forecasts	52	0	600
PF_4 period forecasts on perfect prebaratics	68	63	600

Table XI

Summary of various statistics relating to forecast of precipitation in the period forecasts PF_2 to PF_4 (per cent.)

	a	c	c'	I	J
PF_2	67	77	58	67	17
PF_3	60	62	58	60	0
PF_4	71	85	57	71	27

Table XII

Average and r.m.s. temperature errors (°F) and percentage success in the forecasts F_1 to F_5

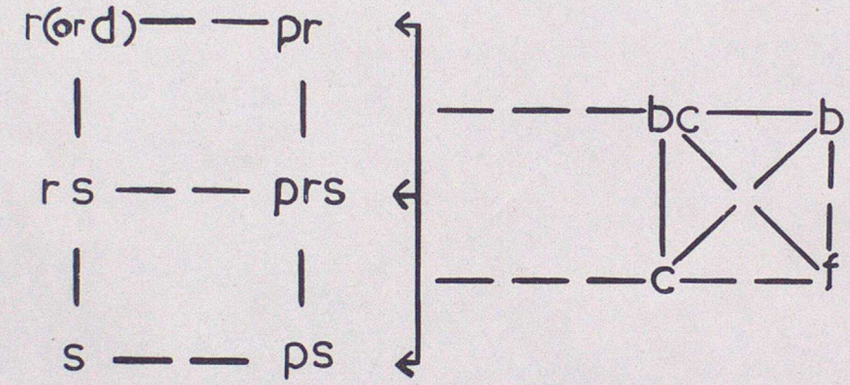
	Average error	r.m.s. error	Percentage success	No. of obs.
F_1 press forecasts on official prebaratics	2.98	3.93	55	528
F_2 control forecasts on official prebaratics	3.14	4.19	45	600
F_3 persistence forecasts	3.55	4.70	0	600
F_4 forecasts on perfect prebaratics	2.82	3.78	59	600
F_5 probable perfect prebaratic forecasts under normal working conditions	2.66	3.49	67	-

Table XIII

Percentage of temperature errors in three ranges

Forecasts	$\leqslant 4°F$	5 to 9°F	$\geqslant 10°F$
F_1	80.7	16.7	2.6
F_2	76.4	20.3	3.3
F_3	72.3	22.5	5.2
F_4	80.7	17.5	1.8
F_5	85.0	14.0	1.0

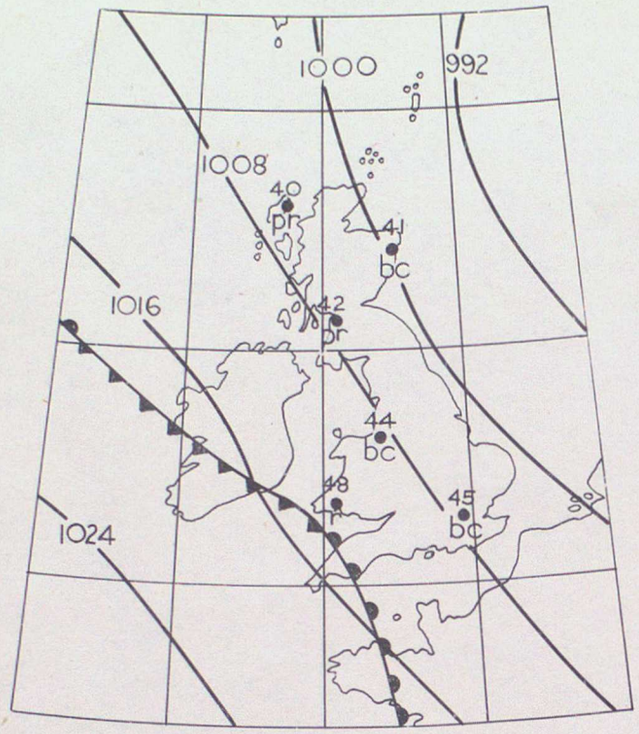
Fig.1 Basis of marking system



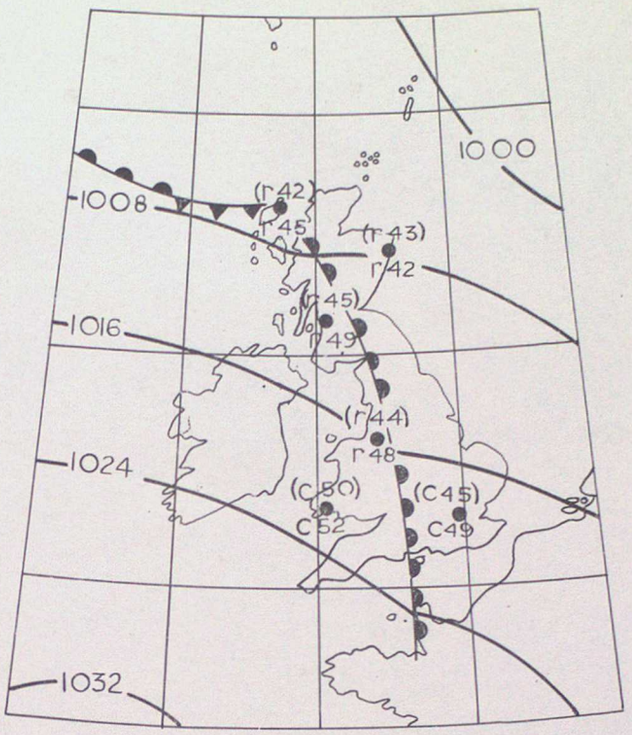
Points to be deducted from the maximum score of 5 are equal to the minimum number of steps separating forecast and observed weather

Fig.2 Weather forecasts and prebaratics for 1200 G.M.T 19:1:51

(a)



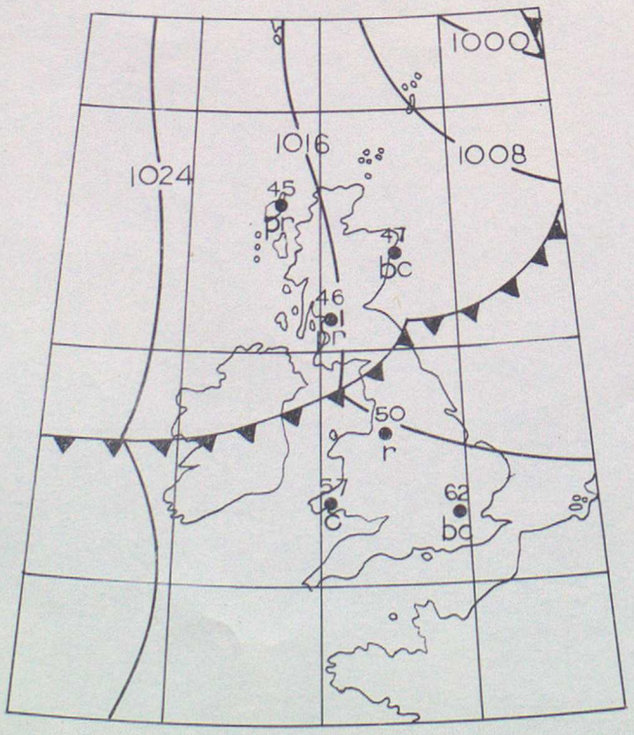
(b)



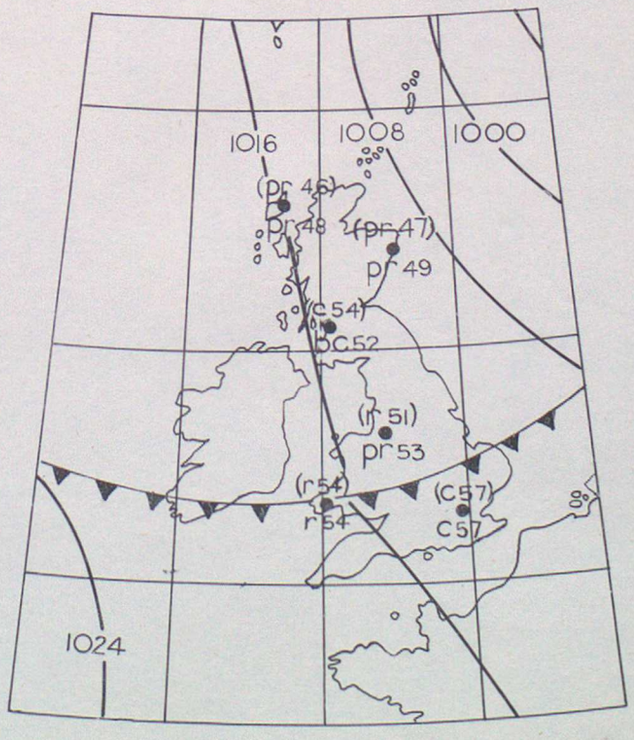
(a) Official prebaratic and press forecasts (b) Perfect prebaratic and test forecasts (observed weather and temperatures in brackets)

Fig.3 Weather forecasts and prebaratics for 1200 G.M.T. 17:9:52

(a)



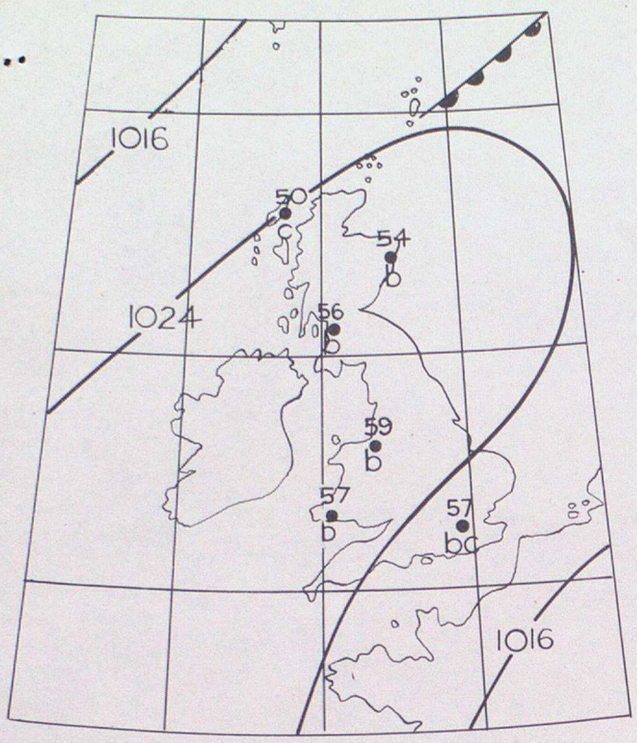
(b)



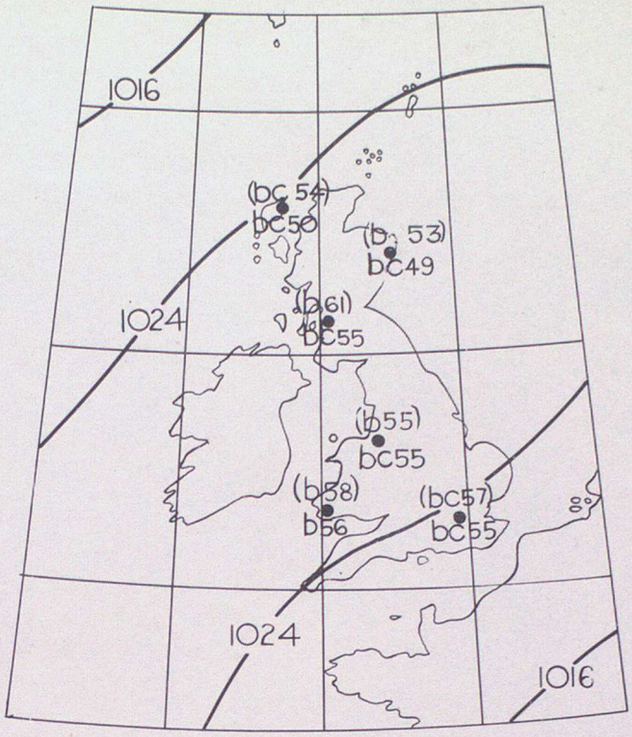
Legend as Fig 2

Fig. 4 Weather forecasts and prebaratics for 1200 G.M.T. 11:5:51

(a)



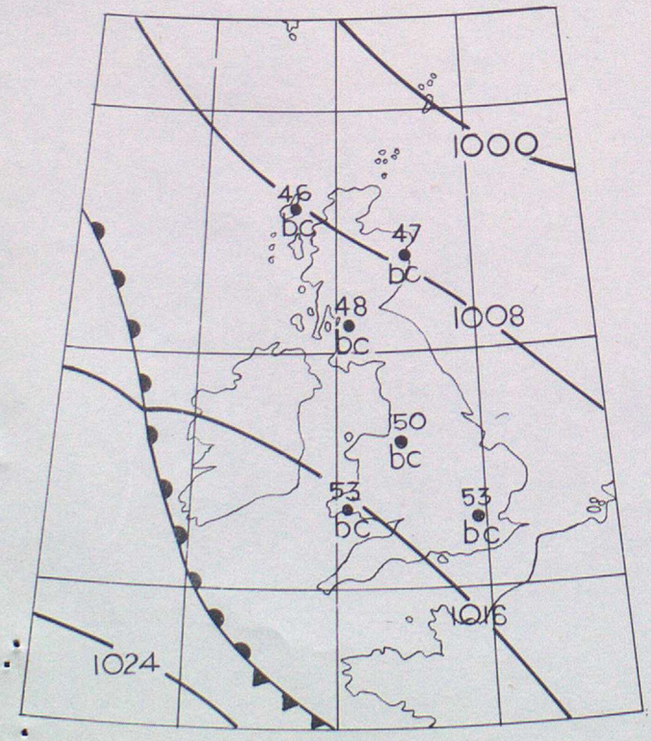
(b)



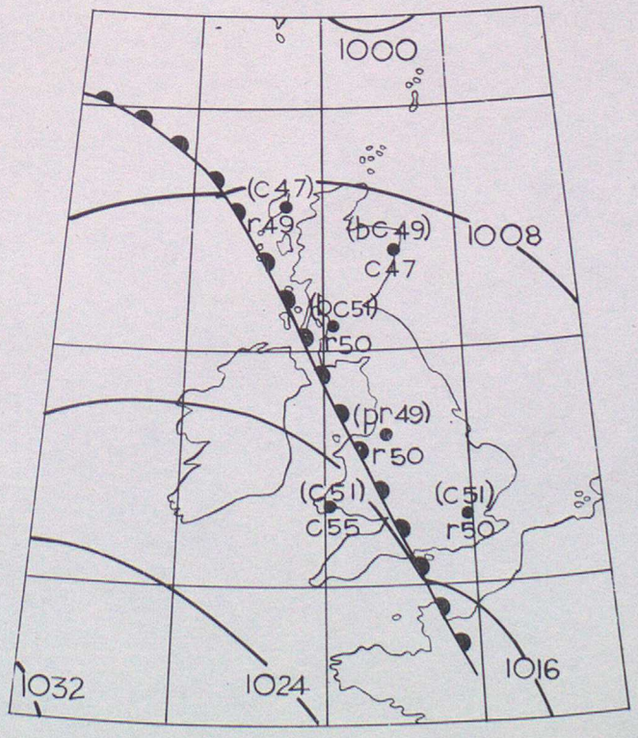
Legend as Fig. 2

Fig. 5 Weather forecasts and prebaratics for 1200 G.M.T. 31:10:52

(a)



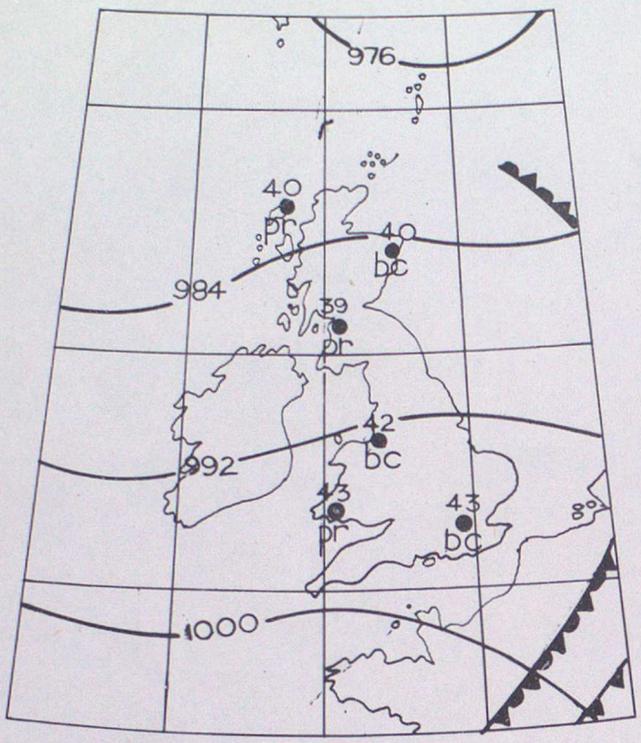
(b)



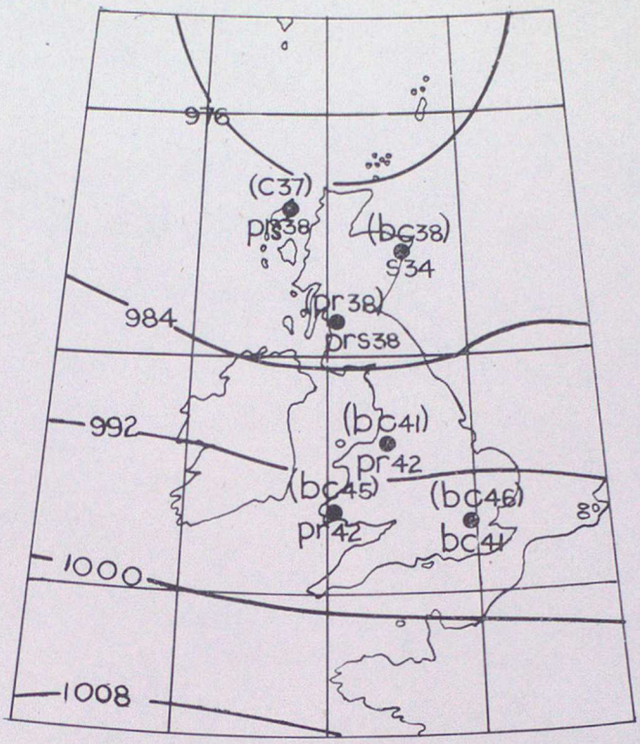
Legend as Fig. 2

Fig. 6 Weather forecasts and prebaratics for 1200 G.M.T. 19:2:51

(a)



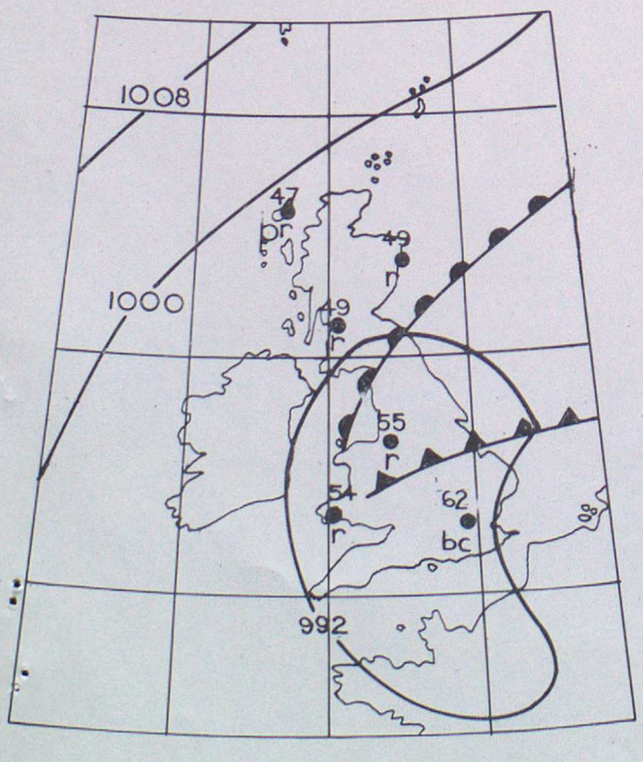
(b)



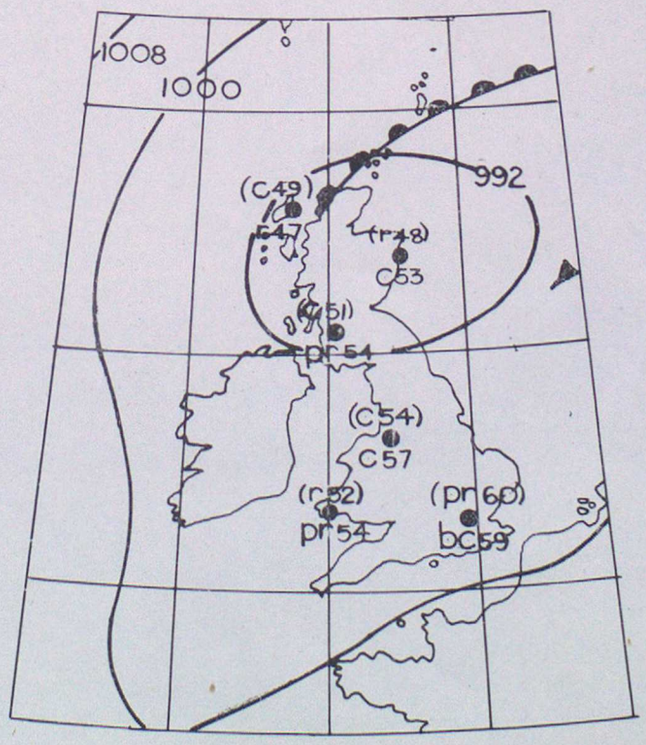
Legend as Fig. 2

Fig. 7 Weather forecasts and prebaratics for 1200 G.M.T. 5:5:52

(a)



(b)



Legend as Fig. 2